



Universität Potsdam

Barbara Krahé, Gün R. Semin

Explaining perceived cross-situational consistency : intuitive psychometrics or semantic mediation?

first published in:
European Journal of Personality, 2 (1988) 4, S. 239-252, DOI
10.1002/per.2410020401

Postprint published at the Institutional Repository of Potsdam University:
In: Postprints der Universität Potsdam
Humanwissenschaftliche Reihe ; 70
<http://opus.kobv.de/ubp/volltexte/2009/3384/>
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-33847>

Postprints der Universität Potsdam
Humanwissenschaftliche Reihe ; 70

Explaining perceived cross-situational consistency: Intuitive psychometrics or semantic mediation?

Gün R. Semin and Barbara Krahé*

University of Sussex, England

Abstract

Recent studies at the interface of social cognition and personality theory have stressed lay persons' ability to 'function as intuitive psychometricians' (Epstein and Teraspulsky, 1986). This research argues that lay persons not only show a substantial degree of accuracy in estimating cross-situational generality of behaviour, but also take into account principles of aggregation over time. In contrast, it is argued here that lay persons' perceptions of the degree of relatedness of different behaviours are mediated largely by the decontextualized semantic relationships between behavioural descriptions. This argument finds support in two experimental studies which demonstrate that the main source for subjects' judgments of 'cross-situational consistency' can be found in an abstracted knowledge base which is represented and mediated through language. The implications of the findings are drawn out for personality research, in particular with reference to domain and item selection in questionnaires for research.

INTRODUCTION

How people's judgments of trait and behaviour covariations are mediated has been at the core of considerable discussion and research since the beginning of the century (Wells, 1907), along with various attempts to account for those conditions where differences between judged and observed covariations are obtained (cf. Cooper, 1981). Mostly, these explanations (e.g., illusory correlations, Chapman, 1967; Chapman and Chapman, 1967; systematic distortion hypothesis, e.g. Shweder, 1982; halo effect, Thorndike, 1920; *inter alia*) have relied on cognitive distortion explanations which are regarded as deriving from 'theories' that raters hold as the principal source to account for biases in covariation judgments.

*Now at the University of Mainz, F.R.G. Requests for reprints should be addressed to: G. R. Semin, The University of Sussex, Psychology Division, Arts D, Falmer, Brighton BN1 9QN, U.K. The research reported here was supported through an ESRC personal grant G 00 24 2033 (Social Psychology of Language and Social Cognition) to the first author and ESRC grant G 00 23 2235.

From the work on covariation detection some authors have readily extrapolated that people's judgments of 'cross-situational consistencies in behavior may be seen as merely another instance of theory-driven covariation assessments operating in the face of contrary evidence' (Nisbett and Ross, 1980, p. 112). However, as Epstein and Teraspulsy (1986) point out, the question of *cross-situational consistency* has not been addressed directly and this extrapolation may in itself be regarded as an overgeneralization. To redress this imbalance between conjecture and evidence, Epstein and Teraspulsy (1986) report a study in which they examined the accuracy of people's perceptions of cross-situational generality in behaviours. The subjects' main task consisted in estimating the degree to which pairs of behaviours were related, given that the behaviours in question had been observed over a period of a month. The items were chosen from studies (i.e. Newcomb, 1929; Mischel and Peake, 1982) where real-life cross-situational correlations were already established. The actual empirical relationships obtained in the original studies correlated significantly with subjects' estimates of these relationships in the Epstein and Teraspulsy (1986) study, suggesting that subjects perform well in their estimates of cross-situational generality of behaviour. Indeed, the important aspect of these findings is that subjects had a reasonably accurate assessment of not only strong positive relationships but also weak relationships between behaviours.

The conclusion that Epstein and Teraspulsy (1986) draw from these results relies on the assumption that people have the ability 'to function as intuitive psychometricians'. This assumption, however, disregards the possibility that other, language-based, factors may be operating in subjects' ratings of predictability. Nevertheless, Epstein and Teraspulsy (1986, p. 1158) conclude, for example, that '. . . subjects functioned as moderately good psychometricians in that they took into account principles of aggregation over occasions and items of behavior.'

However interesting and suggestive Epstein and Teraspulsy's (1986) results may be, the explanation advanced for the findings remains somewhat inchoate. Let us consider what the subjects' task consists of when they are asked to make judgments of relatedness between behavioural items in the abstract (namely, without any reference to a concrete person). Given the task as to whether 'the regularity of class attendance' is related to the 'neatness of appearance', the only source of knowledge that they can consult is the meaning that these sentences have. Thus, when subjects are asked to predict a person's standing on one behavioural item from his or her standing on another item, they would have to do so on the basis of the *semantic meaning* of the items.

A further consideration pertains to how these behavioural items are comprehended across subjects. In the type of task employed by Epstein and Teraspulsy (1986), the meanings of the behaviour items are completely decontextualized and are available only with reference to an abstracted semantic context. Their meanings will therefore be interpreted consensually by subjects, i.e. intersubjective variations will be minimal. Let us briefly illustrate the argument for an abstracted semantic context through contrasting it with meanings in pragmatic contexts (cf. Douglas, 1971). For example, the sentence 'the sun is rising' is a perfectly meaningful sentence in the abstract. However, consider the following two different contexts within which this sentence may have been uttered: (1) Two spies are in the process of bugging a foreign embassy office and this sentence is uttered by one of

them. The meaning of the above sentence changes to 'We better get on with the job, its getting late'. Take another context: (2) A man and a woman are lying in bed, naked, both of them married to somebody else. One of them utters the above sentence. The meaning shifts to 'My husband/wife might be coming soon!', etc. In such pragmatic contexts the sentence acquires an idiosyncratic meaning and has indexical function (Garfinkel, 1967; Garfinkel and Sacks, 1970). In the absence of such pragmatic contexts, however, the sentence can only be understood with reference to an abstract, consensually shared semantic context. One would therefore expect subjects in the experimental situation to draw upon the same knowledge reservoir in making their judgments of predictability and, to the extent that they share the same linguistic community, have consensual interpretations and understandings of them.

On the basis of the above considerations we would like to argue that subjects' predictive judgments are based first and foremost on linguistic conventions pertaining to the meanings of the sentences and on how similar the meaning of one behavioural item is to another in the abstract. Thus, the judgments of subjects about the relatedness of behavioural items is predicated upon semantic mediation rather than based on 'theory driven perceptions' within an intuitive psychometrician's frame of reference. It would therefore follow that if we asked subjects to judge how similar the meaning of one behavioural item is to the meaning of another, we should not only be able to account for Epstein and Teraspulsky's (1986) findings, but also be able to demonstrate that such linguistic factors account for the real-life cross-situational correlations obtained in the Newcomb (1920) and Mischel and Peake (1982) studies in the same way that Epstein and Teraspulsky's (1986) findings do. These propositions were addressed in the two empirical studies reported in this paper.

STUDY 1

Based on the above-outlined rationale, a study was designed to test the following hypotheses:

- (1) Epstein and Teraspulsky's (1986) findings on people's intuitive understanding of cross-situational consistency are a function of the similarity in meaning between pairs of statements which describe trait-related behaviours rather than being driven by an 'intuitive theory of psychometrics'. Ratings of similarity of meanings between pairs of items should therefore give rise to the same empirical patterns of relationships between items as those obtained by Epstein and Teraspulsky on the basis of judgments of predictability of one behaviour item from another. From this it also follows that:
- (2) Ratings of semantic similarity should correspond to the external criteria of behavioural consistency in the same way as Epstein and Teraspulsky's findings do.

In the Epstein and Teraspulsky (1986) study, the trait domains of conscientiousness, friendliness, and extraversion were selected to examine subjects' perceptions

of cross-situational consistency. In each of these domains, seven pairs of behaviour statements were presented covering the range from weak ($r = -0.01$) to strong ($r = 0.73$) empirical relationships previously established for conscientiousness and friendliness by Mischel and Peake (1982) and for extraversion by Newcomb (1929). In order to test the above propositions, the present study employed the same materials used in the Epstein and Teraspulsky study.

Method

Subjects

Sixty first-year undergraduates at the University of Sussex participated in this study on an unpaid voluntary basis. They were students enrolled in different courses in the School of Social Sciences and had no prior knowledge of psychometric theory. All participants were native speakers of English.

Procedure

Subjects were presented with a questionnaire which contained the 21 pairs of behavioural statements (seven in each of the three trait domains) employed by Epstein and Teraspulsky (1986, p. 1154). Among these were pairs of statements such as 'neatness with which the desk is kept/thoroughness with which class notes are taken' (conscientiousness), 'number of people conversed with at breakfast/amount of time spent socializing with friends' (friendliness), and 'resists criticism/high activity level during quiet hour' (extraversion). However, instead of being asked, as in Epstein and Teraspulsky, to estimate the degree of relationship between these behaviour pairs in terms of the predictability of one behaviour on the basis of the other, our subjects' task consisted in '. . . judging how similar the two statements are in terms of their meanings'. Semantic similarity ratings were obtained on a five-point scale ranging from (1) 'not at all similar in meaning' over (3) 'moderately similar in meaning' to (5) 'completely similar in meaning'. After judging semantic similarity for each item-pair, subjects were also asked to indicate their confidence in their judgment on a five-point scale ranging from (1) 'not at all confident' over (3) 'moderately confident' to (5) 'completely confident'. This measure was identical to the one employed by Epstein and Teraspulsky. For details about the stimulus material and its presentation see Epstein and Teraspulsky (1986, p. 1154).

Results

To examine the first hypothesis about whether or not semantic similarity judgments would reproduce the same pattern of results obtained by Epstein and Teraspulsky's perceived consistency ratings, three correlational indices were calculated within each trait domain:

- (a) Firstly, rank-order correlations were computed separately for each subject across the seven behaviour pairs between the semantic similarity judgments and the ratings of perceived consistency established by Epstein and Teraspulsky. These individual correlations were then z-

transformed and averaged across subjects to provide an index of *mean individual correlation* between semantic similarity and perceived consistency.

- (b) Secondly, semantic similarity ratings for each behaviour pair were first averaged across subjects and then correlated with the perceived consistency ratings obtained by Epstein and Teraspulsky. This index will be referred to as *group correlation*.
- (c) Finally, each subject's ratings of semantic similarity were correlated with those of the remaining subjects across the seven behaviour pairs. This index is comparable to the familiar index of corrected item-total correlation used in test construction and expresses the degree to which an individual's score corresponds to that of the sample as a whole. Individual correlations with the group were *z*-transformed and then averaged to provide an index of *mean intra-individual correlation against group*.

Data concerning the confidence ratings for each semantic similarity judgment will not be reported in this section. The rationale for including the confidence ratings in the present study will be more fully understood in connection with the second study reported below, where results on this measure will be presented.

The findings from the correlational analyses are presented in Table 1.

Table 1. Correlations between semantic similarity and perceived consistency (Epstein and Teraspulsky, 1986)

Trait domain	Mean individual correlation	Group correlation	Mean intra-individual correlation against group
Conscientiousness	0.61	0.86	0.62
Friendliness	0.57	0.89	0.61
Extraversion	0.47	0.59	0.65

These results show that semantic similarity judgments correlate substantially with the data on perceived consistency obtained by Epstein and Teraspulsky. Thus, they provide support for the hypothesis that intuitive psychological inferences about the predictability of behaviour are mediated to a significant extent by semantic factors.

Moreover, as will be obvious from Table 2, the pattern of correlations between semantic similarity and perceived consistency is highly similar, indeed almost identical, to the correlations obtained by Epstein and Teraspulsky between perceived and empirical consistency.

The second hypothesis concerning the correspondence between semantic similarity judgments and empirical consistency was examined through a second set of correlational analyses. Using the same correlational indices as described above, semantic similarity judgments were correlated with the original empirical relationships between the behaviour pairs within each trait domain. The results of these analyses are presented in Table 2, which also includes the Epstein and Teraspulsky correlations between perceived and empirical consistency.

Table 2. Correlations between semantic similarity and empirical consistency*

Trait domain	Mean individual correlation		Group correlation		Mean intra-individual correlation against group (identical to Table 1)	
Conscientiousness†	0.61	(0.63)	0.86	(0.96)	0.62	(0.64)
Friendliness†	0.42	(0.58)	0.68	(0.82)	0.61	(0.66)
Extraversion‡	0.34	(0.40)	0.61	(0.57)	0.65	(0.64)

*Epstein and Teraspulsky's correlations between perceived and empirical relationships are given in parentheses.

†Mischel and Peake(1982).

‡Newcomb (1929).

• It can be seen that although the correlations between semantic similarity and empirical consistency are slightly lower on most indices than the figures reported by Epstein and Teraspulsky, the overall pattern is again very similar to the findings of their original study. These results thus provide support for the hypothesis that judgments of semantic similarity account equally well for the empirically observed relationships as the intuitive perception of cross-situational consistencies.

Discussion

The results of the first study support the two hypotheses that were advanced to account for the factors underlying judgments of cross-situational consistency. The findings show that judgments of semantic similarity between pairs of behavioural statements are sufficient to reproduce Epstein and Teraspulsky's (1986) findings which were purported to reflect individuals accuracy in judging cross-situational consistency or behavioural predictability. This suggests that when asked to assess the strength of relationships between behavioural statements presented in an abstract, decontextualized form, individuals refer to linguistic conventions rather than to intuitive notions of personality principles as a source of reference. In doing so, they show a substantial agreement with external empirical criteria. The consensus between individuals as expressed in both the *group correlations* and the *intra-individual against group* correlations further corroborates the interpretation of the present findings in term of semantic mediation.

However, it is possible to argue that judgments of semantic similarity disguise the operation of *theory driven intuitive psychometrics*. It may have been the case that when subjects were giving judgments of semantic similarity, they were in fact drawing upon *intuitive conceptions of cross-situational consistency*, that is, aggregating observations over time. This could imply that what appeared to be a result of linguistic conventions may in fact be reflecting *perceived consistency* due to an implicit reinterpretation of the task by the participants.

STUDY 2

To clarify this ambiguity with respect to the findings of Study 1, a second study was devised to rule out the possibility that ratings of semantic similarity were

confounded with implicit psychological reasoning and to provide a critical test between the two competing explanations.

In the Epstein and Teraspulsky study, subjects were informed that the behaviours described in the pairs of statements had been observed over a period of 1 month. In order to test the assumption that persons are not only capable of making accurate ratings of consistency but are also aware of the principle of aggregation over occasions, these authors also asked their subjects to re-rate two items from each of the conscientiousness and extraversion domains. In this condition, subjects were told that the items were based on observations of only 1 day. Comparing these judgments with their earlier findings, the authors found that relationships were judged to be significantly stronger when behaviours were observed for 30 days than when they were observed for 1 day.

While Epstein and Teraspulsky (1986) interpret these findings as evidence in favour of their subjects' awareness of psychometric principles, we would argue that just the opposite is true. Assuming that observations of behaviour pairs over 30 days must of necessity yield stronger (not more reliable) relationships between the pairs than observations on just 1 day is clearly at odds with a proper understanding of psychometric principles.

A second set of data reported by Epstein and Teraspulsky is more pertinent to the issue of whether naive subjects have some awareness of the principle of aggregation and its relationship to reliability. When they asked their subjects to judge how good a measure of an underlying trait each of a number of behavioural items was, given that they were based on 1 day vs. 1 month of observation, it was found that subjects thought behavioural measures based on 1 month of observation to be better measures of the respective trait than measures observed for only 1 day. However, these data reflecting some awareness by the subjects of the link between aggregation and reliability were not related in any way to the original perceptions of perceived consistency and thus present no evidence against a more parsimonious semantic similarity explanation.

In the present study, this issue was addressed directly and applied to the potential critique that subjects' ratings of semantic similarity in Study 1 may have been influenced or confounded by psychological reasoning. Accordingly, two competing hypotheses were examined:

- (1) If ratings of semantic similarity are confounded or mediated by implicit considerations of consistency over time, then these ratings should be affected by differential information about the reliability of the behavioural items, i.e. subjects should be both more accurate and more confident in their judgments as the number of observations increases.
- (2) If, on the other hand, ratings of semantic similarity are made exclusively on the basis of linguistic conventions, as we would argue, then they should remain unaffected by reliability information, i.e. differential information about the number of observations on which the statements are based should not have an effect on either ratings of semantic similarity or on the confidence with which these judgments are made.

To examine these contrasting hypotheses about semantic mediation vs. perceived cross-situational consistency, subjects in this study were divided into two groups.

One group was told that the behavioural pairs whose semantic similarity they were asked to rate had been derived from 5 days of observation, while the other group was informed that the behavioural statements had been derived from 30 days of observation. Except for this manipulation, the present study consisted in a direct replication of Study 1.

Method

Subjects

A new sample of 60 first-year students in the School of Social Sciences at the University of Sussex participated in this study on an unpaid voluntary basis. As in Study 1, only students who were native speakers of English were included in the sample.

Procedure

Two revised versions of the questionnaire used in Study 1 were created, each of which was randomly distributed to 30 subjects. After the general instructions about the task, subjects in Condition 1 received the following information: 'To provide you with some background information we should briefly explain how these statements were generated: Each pair of statements was obtained by observing a person over a period of 5 days'. Subjects in Condition 2 were informed that 'each pair of statements was obtained by observing a person over a period of 30 days'.¹ Otherwise, the questionnaires were identical to those in Study 1.

Ratings of semantic similarity and confidence were obtained for each behaviour pair on five-point scales (cf. Study 1).

Results

In the first part of the data analysis, judgments of semantic similarity in the two conditions were correlated with Epstein and Teraspulsy's (1986) ratings of perceived consistency. The calculation of the correlation indices was identical to that in Study 1. The findings from these analyses are presented in Table 3.

The overall pattern of correlations between semantic similarity judgments and Epstein and Teraspulsy's perceived consistency data is highly similar to that in the first study. Comparing the effects of observation base, it is evident that correlations are slightly lower under the *30 days* condition as compared to the *5 days* condition. In order to carry out more stringent tests, the following strategy was employed: A random sample of 30 subjects was drawn from the subject pool of Study 1 and added to the two groups of the present study. Within each trait domain, one-way

¹ Epstein and Teraspulsy contrasted subjects' estimates of relationship based on 1 day vs. 1 month of observation. They could only do so because they confined themselves to a subset of two items each from the conscientiousness and extraversion domains. None of these four behaviour pairs happened to refer to the *regularity* of the behaviour in question. In the present study, however, the aim was to examine the impact of differential observation bases on ratings of semantic similarity for the full set of 21 items, many of which refer either explicitly or implicitly to regularities in behaviour (e.g. 'regularity of class attendance'). For these items, information that they had been derived from just 1 day of observation would have been meaningless. Therefore, observation bases of 5 vs. 30 days were used in the present study.

Table 3. Correlations between semantic similarity (observation base: 5 vs. 30 days) and perceived consistency (Epstein and Teraspulsy, 1986)

Trait domain	Mean individual correlation	Group correlation	Mean intra-individual correlation against group
Conscientiousness			
5 days	0.68	0.86	0.71
30 days	0.58	0.85	0.69
(Study 1)	(0.61)	(0.86)	(0.62)
Friendliness			
5 days	0.65	0.94	0.67
30 days	0.48	0.89	0.57
(Study 1)	(0.57)	(0.89)	(0.61)
Extraversion			
5 days	0.63	0.86	0.72
30 days	0.61	0.86	0.68
(Study 1)	(0.47)	(0.59)	(0.65)

analyses of variance were computed for this combined sample of $N = 90$ to examine the effects of observation base (no information, 5 days, 30 days) on mean individual correlations and mean intra-individual correlations. No significant differences were found ($F < 1$).

These findings suggest that neither the correlations between semantic similarity judgments and perceived strength of relationship nor the extent to which each individual's judgments of semantic similarity corresponded to the rest of the group were affected by the manipulation of information about observation base.

As Table 4 shows, a very similar pattern of results was obtained with regard to the correlations between semantic similarity and empirical consistencies.

It is evident that the correlations obtained between semantic similarity and empirical consistency closely correspond to both the results of Study 1 and Epstein and Teraspulsy's figures. Using the procedure described above, one-way analyses of variance were carried out to examine whether the obtained differences between correlations as a function of observation base were significant. Neither the mean individual correlations nor the mean intra-individual correlations against group were significantly different in any of the three trait domains ($F < 1$).

Finally, we examined the hypothesis derived from the *intuitive psychometrician model* that subjects should be more confident in their judgments as the number of observations underlying the behavioural statements increases. The samples of Studies 1 and 2 were combined for this purpose. One-way analyses of variance were conducted to test the effect of observation base (no information vs. 5 days vs. 30 days) on mean confidence ratings across the seven items in each trait domain. The results from this analysis can be seen in Table 5.

Subjects showed a fairly high degree of confidence in their judgments of semantic similarity. More importantly, confidence was unrelated to the number of observations that the subjects believed to have formed the basis of the behavioural statements.

Table 4. Correlations between semantic similarity (observation base: 5 vs. 30 days) and empirical consistency

Trait domain	Mean individual correlation	Group correlation	Mean intra-individual correlation against group (identical to Table 3)
Conscientiousness*			
5 days	0.64	0.86	0.71
30 days	0.58	0.83	0.69
(Study 1)	(0.61)	(0.86)	(0.62)
(Epst/Terasp.)	(0.63)	(0.96)	(0.64)
Friendliness*			
5 days	0.57	0.79	0.67
30 days	0.40	0.68	0.57
(Study 1)	(0.42)	(0.68)	(0.66)
(Epst/Terasp.)	(0.58)	(0.82)	(0.66)
Extraversion†			
5 Days	0.43	0.71	0.72
30 Days	0.45	0.46	0.68
(Study 1)	(0.34)	(0.61)	(0.65)
(Epst/Terasp.)	(0.40)	(0.57)	(0.64)

*Mischel and Peake (1982).

†Newcomb (1929).

Table 5. Mean confidence ratings as a function of observation base

Trait domain	Observation base			
	Study 1	Study 2		
		No Info.	5 days	
Conscientiousness	3.84	3.80	3.81	($F < 1$, n.s.)
Friendliness	3.82	3.82	3.85	($F < 1$, n.s.)
Extraversion	3.80	3.93	4.01	($F = 1.91$, n.s.)

Discussion

The present results strongly confirm the findings from the first study and consolidate further their interpretation in terms of semantic mediation. Evidence suggests that when persons are asked to judge the semantic similarity between pairs of behavioural statements, they approach the task as competent users of their native language and its conventions rather than as intuitive psychometricians. In drawing upon their semantic knowledge, individuals show a high degree of confidence as well as substantial intersubjective agreement which is unaffected by *psychometrically relevant* information.

As far as the subjects' awareness of the principle of aggregation is concerned, findings from the Epstein and Teraspulsy study were inconclusive. While their subjects did recognize that multiple observations of behaviours provide better measures of the underlying trait than single measures, they failed to apply this knowledge to the judgment of perceived consistency and erroneously assumed a linear positive relationship between period of observation and level of consistency.

The findings of the present study suggest that—notwithstanding their potential awareness, in principle, of the importance of aggregation—there was no need for our subjects to draw upon this knowledge in order to reproduce not only the patterns of perceived consistency obtained by Epstein and Teraspulsy but also empirically obtained behavioural consistencies. Information about observation base is simply not relevant to the type of judgments people make when asked to rate the relatedness of decontextualized behavioural descriptions and, as a consequence, leaves both the accuracy and the confidence of such ratings unaffected.

GENERAL DISCUSSION

The findings of both studies jointly suggest that the main source for subjects' judgments of 'cross-situational consistency' is to be found in an abstracted knowledge basis which is represented in and mediated by language. In itself these findings question the idea of an *implicit* theory, be it of a psychometric nature or otherwise. In fact, it may be misleading to suggest a multitude of different implicit theories as they proliferate currently for distinct domains of social and individual behaviour.² It may be argued that semantic conventions provide the *general source* for everyday theories about a variety of domains. In fact, it would be surprising if language did not constitute the major source for such theories, since it is the main reservoir into which our cultural experience is deposited. Consequently, one can expect that studies such as the present ones, which utilize one specific method of evoking semantic relationships, reveal an alternative explanation for Epstein and Teraspulsy's findings.

On the basis of the above considerations, one could argue that advancing the idea of an 'intuitive psychometrician' in order to account for the results obtained by Epstein and Teraspulsy involves what one may call a *psychomorphism*. It would appear to be the case that linguistic conventions constitute the source to which subjects' predictability judgments can be traced. That subjects are able to answer questions worded in terms of predictability does not mean, however, that in being able to do so they are acting as *intuitive psychometricians*. We therefore think that the attribution of a specific *implicit or intuitive theory* to subjects (i.e. psychomorphism) is unnecessary.

The present findings in conjunction with Epstein and Teraspulsy's (1986) results raise another issue to do with judgments about behaviour covariations given in the abstract and actual behaviour covariations observed over time. The two sets of results, namely judgments about predictability as well as similarity in meaning, both given in the abstract, correlate satisfactorily with empirically obtained criteria for behavioural patterns observed in the original studies of Newcomb (1929) and Mischel and Peake (1982). These findings would imply that there is a *correspondent* relationship between linguistic conventions and actual behaviour. In fact, both sets of findings, but particularly the present ones, would appear to directly contradict a popular theory driven hypothesis, namely the systematic distortion hypothesis

² Indeed, there is evidence suggesting that such theories may be quite explicit, at least in the personality domain (cf. Semin and Krahe, 1987; Semin, Rosch and Chassein, 1981; Semin and Chassein, 1985) and possibly even in the intelligence domain (e.g. Sternberg, Conway, Ketron and Bernstein, 1981).

(SDH) (Shweder and D'Andrade, 1980), which maintains that 'inferences about personality contain a systematic bias in that propositions about "what is like what" are substituted for propositions about what is likely, and *memory* for personality relevant events contains a systematic bias in that attitudes, affects, and behaviors that are conceptually associated . . . are recalled as if they covaried' (Shweder, 1982, p. 66). Indeed, the contradiction is even more apparent if one considers that one of the empirical studies examining the SDH (Shweder, 1975) utilizes (among others) the empirical data on extraversion-introversion in Newcomb's (1929) study. Although the overall pattern of relationships reported by Shweder does not seem to support the SDH, a re-analysis of a subset of the data (33 out of 110 correlations) produces a result which appears to be supportive of this hypothesis. It should be noted, however, that there are severe conceptual and methodological problems with this particular study (Block, Weiss and Thorne, 1979) and the SDH in general (Semin and Greenslade, 1985; *inter alia*). On the basis of our studies, we would like to suggest that, contrary to the SDH, abstracted judgments of 'what is like what' actually correspond to 'what goes with what' in empirical behaviour covariations. The question, however, remains why this should be the case. A possible answer is suggested in the following reasoning.

In any study on behavioural consistency, the first step consists in a selection of behaviours that belong to a specific domain under investigation. Indices of consistency are then obtained through examining the covariation between behavioural criteria within the specific domain. The issue of how these behavioural criteria are selected is generally not addressed explicitly in studies of cross-situational consistency. The question that arises in this context is: What are the resources that an investigator can draw upon in order to generate instances of behaviour which may then be utilized as indicators of consistency? On the basis of the present findings, it would seem plausible to argue that such a selection must be largely guided by an intuitive understanding, mediated by linguistic conventions, according to which a set of behaviour instances is perceived to belong to the same domain. For example, what basis is there to assume that the 'neatness with which the bed is made' taps the same behavioural domain (namely, conscientiousness) as 'regularity of class attendance'? From the above considerations we would like to conclude that language furnishes the general knowledge reservoir from which not only judgments of predictability and similarity in meaning are derived, but also empirical criteria for behavioural consistency are selected. This argument allows us to explain the recurrent patterns of findings observed over actual behaviour covariations, judgments of predictability, and judgments of similarity in meaning. It would seem well advised to pay more attention to the role played by language and linguistic conventions as they enter both intuitive and professional theorizing about personality. We would therefore like to suggest that the time has come to appreciate more fully the capacity of both lay persons and psychologists as competent members of their language communities.

REFERENCES

- Block, J., Weiss, D. S. and Thorne, A. (1979). 'How relevant is the semantic similarity interpretation of personality ratings?', *Journal of Personality and Social Psychology*, **37**: 1055-1074.

- Chapman, L. J. (1967). 'Illusory correlation in observational report', *Journal of Verbal Learning and Verbal Behavior*, **6**: 151-155.
- Chapman, L. J. and Chapman, J. P. (1967). 'Genesis of popular but erroneous psychodiagnostic observations', *Journal of Abnormal Psychology*, **72**: 193-204.
- Cooper, W. H. (1981). 'Conceptual similarity as a source of illusory halo in job performance ratings', *Journal of Applied Psychology*, **66**: 302-307.
- Douglas, J. (1971). *Understanding Everyday Life*, Routledge & Kegan Paul, London.
- Epstein, S. and Teraspulsky, L. (1986). 'The perception of cross-situational consistency', *Journal of Personality and Social Psychology*, **50**: 1152-1160.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*, Prentice Hall, New York.
- Garfinkel, H. and Sacks, H. (1970). 'The formal properties of practical actions'. In: McKinney, J. C. and Tiryakian, E. A. (Eds), *Theoretical Sociology*, pp. 124-178, Doubleday: Anchor Books, New York.
- Mischel, W. and Peake, P. K. (1982). 'Beyond *déjà vu* in the search for cross-situational consistency', *Psychological Review*, **89**: 730-755.
- Newcomb, T. N. (1929). Consistency of certain extrovert-introvert behavior patterns in 51 problem boys. *Contributions to Education*, No. 382, Columbia University, Teachers College, New York.
- Nisbett, R. E. and Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, NJ.
- Semin, G. R. and Chassein, J. (1985). 'The relationship between higher order models and everyday conceptions of personality', *European Journal of Social Psychology*, **15**: 1-15.
- Semin, G. R. and Greenslade, L. (1985). 'Differential contributions of linguistic factors to memory-based ratings: systematizing the systematic distortion hypothesis', *Journal of Personality and Social Psychology*, **49**: 1713-1723.
- Semin, G. R. and Krahe, B. (1987). 'Lay conceptions of personality: Eliciting tiers of a scientific conception of personality', *European Journal of Social Psychology*, **17**: 199-209.
- Semin, G. R., Rosch, E. and Chassein, J. (1981). 'A comparison of the common sense and "scientific" conceptions of extraversion-introversion', *European Journal of Social Psychology*, **11**: 77-86.
- Shweder, R. A. (1975). 'How relevant is the individual difference theory of personality?', *Journal of Personality*, **43**: 455-484.
- Shweder, R. A. (1982). 'Fact and artifact in trait perception: The systematic distortion hypothesis'. In: Maher, B. A. and Maher, W. B. (Eds), *Progress in Experimental Personality Research*, Vol. 11, pp. 65-101, Academic Press, New York.
- Shweder, R. A. and D'Andrade, R. G. (1980). 'The systematic distortion hypothesis'. In: Shweder, R. A. (Ed), *Fallible Judgment in Behavioral Research. New Directions for Methodology of Social and Behavioral Science*, pp. 3-38, Jossey-Bass, San Francisco.
- Sternberg, R. J., Conway, B. E., Ketron, J. L. and Bernstein, M. (1981). 'People's conceptions of intelligence', *Journal of Personality and Social Psychology*, **41**: 37-55.
- Thorndike, E. L. (1920). 'A constant error in psychological ratings', *Journal of Applied Psychology*, **4**: 25-29.
- Wells, F. L. (1907). 'A statistical study of literary merit', *Archives of Psychology*, **1** (7).

RÉSUMÉ

De récentes recherches sur la séparation entre cognition sociale et théorie de la personnalité ont souligné la faculté des profanes de fonctionner comme des 'psychométriciens intuitifs' (Epstein and Teraspulsky, 1986). Ces recherches montrent que les profanes font non seulement preuve d'un haut niveau de précision dans l'évaluation de la généralité transsituationnelle du comportement, mais encore qu'ils tiennent compte des principes de l'agrégation sur le temps. Dans cet article, par contre, on pose que les mesures des relations entre les différents comportements observées par des profanes sont fortement médiatisées par les relations sémantiques décontextualisées entre les descriptions comportementales. Cette argumentation est étayée par deux recherches expérimentales qui montrent que la source principale du jugement de la 'consistance transsituationnelle' peut être trouvée dans

la connaissance abstraite qui est représentée et médiatisée par le langage. Les implications de ces résultats pour la recherche en personnalité sont exposées; en référant particulièrement au choix du domaine et à la sélection de l'item pour la construction de questionnaires de recherche.

ZUSAMMENFASSUNG

Neuere Untersuchungen an der Schnittstelle von Persönlichkeitspsychologie und sozialer Kognition heben die Fähigkeit von Laien hervor, über 'intuitives psychometrisches Wissen' zu verfügen (Epstein and Teraspulsky, 1986). In diesen Arbeiten wird argumentiert, daß Laien in ihren Einschätzungen der situationsübergreifenden Konsistenz von Verhalten nicht nur ein beträchtliches Maß an Genauigkeit zeigen, sondern bei ihren Urteilen auch das Prinzip der Aggregierung von Beobachtungen über mehrere Zeitpunkte berücksichtigen. Im Unterschied hierzu wird in der vorliegenden Arbeit die These vertreten, daß die Urteile von Laien über die Zusammengehörigkeit unterschiedlicher Verhaltensweisen weitgehend von den dekontextualisierten semantischen Beziehungen zwischen den Verhaltensbeschreibungen bestimmt werden. Diese These findet in zwei Experimenten Bestätigung, in denen gezeigt wird, daß Urteile naiver Vpn über die 'situationsübergreifende Konsistenz' von Verhalten unter Rückgriff auf abstrakte Wissensbestände gefällt werden, die sprachlich vermittelt und repräsentiert werden. Abschließend werden die Konsequenzen der Ergebnisse für die Persönlichkeitsforschung, insbesondere bezogen auf die Auswahl von Verhaltensbereichen und Items in der persönlichkeits-psychologischen Forschung, diskutiert.